Second UK Mplus User Group Meeting

The University of Manchester

Tuesday 15th March – Wednesday 16th March 2011

Manchester Dental Education Centre

All presentation sessions will be held in the lecture theatre at MANDEC

All tea and coffee breaks and lunch will be served in the atrium area at MANDEC

Organising Committee:

Dr Tim Croudace, University of Cambridge

Dr Jon Heron, University of Bristol

Dr Richard Emsley, University of Manchester

Second UK Mplus User Group Meeting Programme

Key:
P – 45 minute oral presentation and 10 minutes questions
M – 30 minute oral presentation and 5 minutes questions
S – 15 minute oral presentation and 5 minutes questions

Tuesday 15th March

10.00 – 10.15	Registration and arrival

10.15 – 12.15   Morning Session 1 (Chair: Dr Richard Emsley)

Introduction
Dr Richard Emsley, The University of Manchester

P1: Estimating treatment efficacy from a randomised controlled trial with variable degrees of non-compliance with treatment allocation and subsequent loss to follow-up
Prof Graham Dunn, The University of Manchester

S2: Genetic susceptibility to language development and Attention-Deficit/Hyperactivity Disorder
Dr Tezenas du Montcel, INSERM

S10: Early life patterns of common infection: A latent class analysis
Mrs Sarah Fleming, University of Leeds

S6: Latent class analysis of patterns of occurrence of proteinuria during normal term pregnancy
Mrs Corrie Macdonald-Wallis, University of Bristol

12.15 – 13.00   Lunch

13.00 – 14.30   Afternoon Session 1 (Chair: Dr Jon Heron)

M4: Item response modelling in Mplus: features and applications.
Dr Anna Brown, University of Cambridge

S5: The use of an Exploratory Structural Equation Modelling (ESEM) Framework in Mplus to Guide Instrument Revision: Application to the Family Perceptions Scale (FPS)
Dr Paul Tiffin, Durham University

M3: Longitudinal measurement of salivary cortisol in adolescents: the relationship to depression and the influence of 5HTTLPR and adverse life events
Dr Matthew Owens, University of Cambridge

14.30 – 15.00   Tea/Coffee Break

15.00 – 17.00   Afternoon Session 2 (Chair: Dr Tim Croudace)

S4: Bayesian bifactor analysis in Mplus: an application to data from a British mental health survey
Mrs Rachel McCrea, University College London

S9: Class-based effects of media violence on aggression: A multi level latent growth model approach investigating individual and group differences
Mr Robert Busching, University of Potsdam

P2: Bayes in Mplus
Prof Bengt Muthén, UCLA (live video talk)

17.30 – 19.00   Social event at Krobar on Oxford Road

19.30 – 21.00   Informal dinner at Spicy Hut restaurant in Rusholme

Wednesday 16th March

9.00 – 11.00   Morning Session 1 (Chair: Dr Jon Heron)

S1: Defining Cancer Survivorship using Structural Equation Modelling
Dr Katie Harris, Biostatistics, University of Leeds

S3: Autoregressive latent trajectory models: an empirical example using functional limitation and psychological distress in rheumatoid arthritis
Mr Sam Norton. University of Hertfordshire

M2: Multilevel, path analytic obesogenic model: application to neighbourhood deprivation and individual obesity in the UK
Dr Gindo Tampubolon, University of Manchester

S8: Multiple imputation multivariate multilevel model of well-being across Europe during the crisis
Mr Adi Pierewan, University of Manchester

S7: Modeling the heterogeneity of cognitive decline using univariate and bivariate growth models
Dr Graciela Muniz Terrera, MRC Biostatistics Unit, University of Cambridge

11.00 – 11.30   Tea/Coffee Break

11.30 – 13.00   Morning Session 2 (Chair: Dr Richard Emsley)

M5: Creating and running batches of models within R: Practical introduction to Mplus Automation package
Dr Jan Stochl, University of Cambridge

S11: A latent class approach to adolescent alcohol patterns: the ALSPAC study
Mr Roberto Melotti, University of Bristol

M1: Multilevel finite mixture model for unobserved heterogeneous demand for health care
Mr Sujarwoto Sujarwoto, ISC, University of Manchester

13.00	Mplus Meeting Closes

14.00 – 17.00	**Biostatistics Group Seminar on "Latent Variable Modelling"**
14.00 – 14.50	**Intergenerational correlations in size at birth and the contribution of environmental factors: the Uppsala Birth Cohort Multigenerational Study, Sweden, 1915-2002** Dr Bianca De Stavola, LSHTM
14.50 – 15.40	**Psychometric epidemiology, Rasch psychiatry and advances towards well-being** Dr Tim Croudace, University of Cambridge
15.40 – 16.00	Tea/Coffee Break
16.00 – 16.50	**Interactions, effect moderation and random coefficients in latent variable models: identifying patterns of prenatal vulnerability** Prof Andrew Pickles, Institute of Psychiatry, King's College London

Second UK Mplus User Group Meeting Abstracts

P1: ESTIMATING TREATMENT EFFICACY FROM A RANDOMISED CONTROLLED TRIAL WITH VARIABLE DEGREES OF NON-COMPLIANCE WITH TREATMENT ALLOCATION AND SUBSEQUENT LOSS TO FOLLOW-UP

Prof Graham Dunn, Biostatistics Group, University of Manchester

The ODIN trial (Dowrick *et al.*, 2000; Dunn et al., 2003) evaluated the effect of psychological intervention for the treatment of depression in primary care. Psychotherapy was not available to those allocated to the control condition and, of those allocated to psychotherapy, only 54% received the therapy as planned – full compliance (31% either refused therapy or failed to turn up – noncompliance, and the remaining 15% started but then discontinued their treatment – partial compliance). Failure to provide 6-month outcome data was associated with treatment compliance. In this talk we describe the use of finite mixture modelling (CACE estimation) in *Mplus* to evaluate the effects of treatment in the presence of variable non-compliance (noncompliance, partial and full compliance) and subsequent loss to follow-up. Our main concern is model selection, paying particular attention to the missing data mechanism and associated assumptions. The work builds on that already described in Dunn *et al*. (2005).

References

Dowrick, C.F., Dunn, G**.** *et al*. (2000). Problem-solving treatment and group psycho-education for depression: a multicentre randomized controlled trial. *British Medical Journal* **321**, 1450-1454.

Dunn, G., Maracy, M. *et al*. (2003). Estimating psychological treatment effects from an RCT with both non-compliance and loss to follow-up. *British Journal of Psychiatry* 183, 323-331.

Dunn, G., Maracy, M. & Tomenson, B. (2005). Estimating treatment effects from randomized clinical trials with non-compliance and loss to follow-up: the role of instrumental variable methods. *Statistical Methods in Medical Research* **14**, 369-395.

P2: BAYES IN MPLUS

Prof Bengt Muthén, UCLA

SEE SLIDES ON ADDITIONAL HANDOUT

M1: MULTILEVEL FINITE MIXTURE MODEL FOR UNOBSERVED HETEROGENEOUS DEMAND FOR HEALTH CARE

Mr Sujarwoto Sujarwoto, Institute for Social Care, University of Manchester

A more accurate estimation for health demand care is crucial for better allocation policy and health planning purposes. Inpatient length of stay in hospital (LOS) is well known as a reasonable proxy of resource consumption and health care demand (Leyland and Boddy 1997; Lee et al., 1998). Therefore, accounting for distributional characteristics of LOS is important as prescriptive policies for more efficient utilization of health care resources. However, heterogeneity of LOS within latent groups introduces a problem into the statistical analysis (Marazzi et al., 1998, :915). Typical LOS distributions are skewed (asymmetric) and contain outliers. Both of these features are serious hindrances for the use of the arithmetic mean as an estimate of the expected LOS. Skewness is characterized by a long-sided tail (generally toward high values in the case of LOS); it limits the use of inference techniques based on the normality assumption (e.g., confidence intervals and tests for means).

We use multilevel finite mixture negative binomial models for dealing with heterogeneity of LOS within latent groups of impatient treatments. The heterogeneity of LOS observations is modelled by a two-component mixture model, with one component corresponding to the short-stay sub-population and another component corresponding to the long-stay sub-population. Since these two latent components are nested within regional development, we include random intercept simultaneously in the model. This multilevel component allows us to capture the effect of variation on local policies on health care demand. Deb and Trivedi (1997, :313) affirm that finite mixture negative binomial model is a better framework than the hurdle model in distinguishing between infrequent and frequent user of health care. This model provides a natural representation since each latent class can be seen as a `type' of individual, while still accommodating heterogeneity within each class. MPlus provides suitable tool to run this model.

We apply this model for nationwide data on inpatient healthcare in Indonesia. Associated risk factors for short-stay and long-stay subgroups were identified from the respective negative binomial components. In addition, significant heterogeneities within each group were also found. An encouraging feature of the health care system in Indonesia emerging from the findings of this study is that inpatient care in hospitals is accessible to patients regardless of economic and educational status. This finding is important in view of the near absence of health insurance and the fact that healthcare expenditure is borne by patients. The effect of district and regional development characteristics are also apparent from random intercept model. Inpatient hospital-based healthcare facilities seem to be biased towards urban and more developed districts in terms of concentration of modern facilities and efficiency of treatment. Taking this into account, it is desirable that the public health care system in the country should allocate more on rural and less developed areas.

Keyword: multilevel finite mixture model, demand for health care.

M2: MULTILEVEL, PATH ANALYTIC OBESOGENIC MODEL: APPLICATION TO NEIGHBOURHOOD DEPRIVATION AND INDIVIDUAL OBESITY IN THE UK

Dr Gindo Tampubolon, University of Manchester

While there are growing number of studies examining the association between neighborhood physical and services characteristics and body mass index (BMI), fewer studies have examined the potential influence of neighborhood social environments on an individual's BMI. Using the Welsh Health Survey 2007 and the Living in Wales 2007 survey, we assessed the contribution of neighborhood social capital to explaining variation in obesity (BMI >=30). After adjusting for consumption of fruits and vegetables (caloric intake), regular exercise, gender, age (in 5 year groups), employment status, social class, housing tenure or wealth, and education, we found strong associations between obesity and neighborhood social capital. Living in trustworthy neighbourhoods has the odds ratio of 0.923 (p value = 0.017), living in friendly neighbourhoods increases the chance of being obese with the odds ratio of 1.172 (p value = 0.050). Our study suggests the potential role of social environment in explaining variations in obesity. Results from England surveys will also be discussed.

M3: LONGITUDINAL MEASUREMENT OF SALIVARY CORTISOL IN ADOLESCENTS: THE RELATIONSHIP TO DEPRESSION AND THE INFLUENCE OF 5HTTLPR AND ADVERSE LIFE EVENTS

Dr Matthew Owens, University of Cambridge

M4: ITEM RESPONSE MODELLING IN MPLUS: FEATURES AND APPLICATIONS

Dr Anna Brown, University of Cambridge

M5: CREATING AND RUNNING BATCHES OF MODELS WITHIN R: PRACTICAL INTRODUCTION TO MPLUSAUTOMATION PACKAGE

Dr Jan Stochl, University of Cambridge

S1: DEFINING CANCER SURVIVORSHIP USING STRUCTURAL EQUATION MODELLING.

Dr Katie Harris, Biostatistics, University of Leeds

Background -As cases of cancer continue to rise, and deaths fall, there is a growing population of survivors. To date, cancer survivorship has largely been investigated in the domain of small interview based (qualitative) studies. Our study introduces a novel (quantitative) method for exploring survivorship. Cancer survivorship will be examined by considering the clinical attendance patterns (NHS 'footprints') for populations of cancer survivors, for three distinct cancer sites. The aim is to describe patterns of survivorship and 'types' of cancer survivor using structural equation modelling, undertake in Mplus software.

Methodology - NHS 'footprint' is a concept that is not directly measurable. It may be, however, described as a latent variable by a combination of features that are observed. Variables such as stage of diagnosis, the number, duration and interval between patient's hospital episodes will be considered indicators of NHS 'footprint'. Variables such as reason for hospital admission, treatment type and deprivation are considered explanatory. NHS footprint is then related to patient outcome (alive or dead, secondary cancer or recurrence). An added complexity is that the data form a hierarchical structure, whereby episodes are nested within patients. To exploit the data fully, this hierarchy must be taken into account by modelling patient and episode level NHS 'footprint'.

Discussion - An advantage of representing the data using latent variables is that the dimensionality of the data is reduced. This is necessary as a large number of variables are collected on each subject, which are repeated for each episode. This study demonstrates the utility and application of structural equation modelling to a large complex dataset, with complex hierarchical structure. Furthermore, there is great importance in the application as, assessing differences for different cancer survivors may facilitate identification of a variety of support requirements for cancer survivors, as well as associated adverse events. A future aim is to generalise this methodology for all cancer sites.

S2: GENETIC SUSCEPTIBILITY TO LANGUAGE DEVELOPMENT AND ATTENTION-DEFICIT/ HYPERACTIVITY DISORDER

Dr Tezenas du Montcel, INSERM

Attention-deficit/hyperactivity disorder (ADHD) is the most prevalent psychiatric disorder emerging during childhood and is frequently associated with language deficits. It has been shown that early hyperactivity and/or inattention (H/I) symptoms negatively affect language development, which in turn contribute to H/I symptoms by early school-age. Preschoolers with both elevated levels of H/I and language delays may thus be more at risk of ADHD by school-age. Genetic background has been shown to contribute both to ADHD and to language development. The purpose of the present study was to examine the genetic influence on the developmental course of the association between language skills and H/I between 18 months and 7 years using genetically informative data from the Québec Newborn Twin Study (QNTS). Participants were twins assessed on language skills and H/I symptoms at 18 months, 30 months, 5 years, and 7 years. Several candidate genes were genotyped for both twins and their parents when available.Twin studies are designed to estimate the relative contribution of genes, common environment, and unique environment on the phenotype's variance by comparing within- and across-pair similarity of twin pairs as a function of zygosity. The method is based on the fact that monozygotic twins (MZ) share 100% of their genes and dizygotic twins (DZ) share, on average, 50% of their genes, and that twin pairs are raised in the same family. To test for the influence of candidates genes on the biometric cross-lagged model, we modified the usual twin models in order to take in account the exact percentage of genes shared between the two DZ twins. Results will be shown.

S3: AUTOREGRESSIVE LATENT TRAJECTORY MODELS: AN EMPIRICAL EXAMPLE USING FUNCTIONAL LIMITATION AND PSYCHOLOGICAL DISTRESS IN RHEUMATOID ARTHRITIS

Mr Sam Norton, University of Hertfordshire

Autoregressive latent trajectory (ALT) models combine models of group change (autoregressive) with models of individual change (latent growth) allowing for the synthesis of two approaches often perceived to be competing. An empirical example of a bivariate ALT model using data concerning functional limitation and psychological distress in rheumatoid arthritis will be presented using data from an ongoing prospective inception cohort (N=784). The analysis extends previous research in showing that rather than a cross-lagged association, changes in psychological distress occur synchronously with changes in functional limitation and are associated with the progressive functional decline observed in this condition. The talk will discuss the implementation of ALT models in Mplus including issues that arise when the latent trajectory is non-linear and when covariates are included in the model.

S4: BAYESIAN BIFACTOR ANALYSIS IN MPLUS: AN APPLICATION TO DATA FROM A BRITISH MENTAL HEALTH SURVEY

Mrs Rachel McCrea, University College London

Many measures of mental health, while designed to measure a single construct, are in fact comprised of a number of 'parcels' of related content. For example, a questionnaire designed to measure depression may contain a number of items relating to each of the three broad areas of low mood, somatic symptoms and distorted thinking. When a standard item response theory (IRT) or single factor model is used to model responses to such a questionnaire, these blocks of related content may violate the conditional independence assumption of such models, since there are likely to be residual correlations between items in the same content area.

The bifactor model, which can account for these residual correlations through the addition of uncorrelated specific factors for each content area, is becoming an increasingly popular method for modelling such data. However, the added complexity of these models, particularly when there are many specific factors, can lead to computational difficulties. Bayesian Markov chain Monte Carlo (MCMC) methods can make such complex analyses more tractable.

This presentation will describe some of the practical issues encountered in the use of the new Bayes estimator in Mplus to estimate a bifactor model. The model is applied to data on the symptoms of common mental disorders from a British mental health survey. The talk will consider the following areas: the motivation for using a bifactor model on this particular dataset; advantages of a Bayesian analysis relative to other methods of estimation for the bifactor model; issues related to model identification, choice of prior distributions and convergence of the MCMC chains; and finally, some benefits of using the Bayes estimator in Mplus as opposed to estimating the bifactor model in WinBUGS (a piece of free software for the application of MCMC methods in Bayesian analysis).

S5: THE USE OF AN EXPLORATORY STRUCTURAL EQUATION MODELLING (ESEM) FRAMEWORK IN MPLUS TO GUIDE INSTRUMENT REVISION: APPLICATION TO THE FAMILY PERCEPTIONS SCALE (FPS)

Dr Paul Tiffin, Durham University

Traditionally, measurement models have been validated by a sequence of exploratory then confirmatory factor analyses (EFA and CFA). This can lead to CFA models that require repeated modification in an effort to achieve acceptable fit. Using the recently published Family Perceptions Scale as an example, this study illustrates how Likert scale-based psychometric instrument revision can be guided by utilising the recently described approach of Exploratory Structural Equation Modelling (Asparouhov & Muthén, 2009).

In addition, related challenges related to the modelling of response patterns from ordinal-based rating instruments are discussed. Possible solutions to these issues, implemented in the Mplus software package, are highlighted. The original structure of the FPS is tested using a conventional CFA. Item level data from the original pilot sample (N=673) is then reanalysed using an ESEM approach to guide the restructuring of the instrument in order to achieve simple structure thus improve fit of the CFA model. Study power is confirmed using a Monte Carlo simulation study.

S6: LATENT CLASS ANALYSIS OF PATTERNS OF OCCURRENCE OF PROTEINURIA DURING NORMAL TERM PREGNANCY

Mrs Corrie Macdonald-Wallis, University of Bristol

Authors: Corrie Macdonald-Wallis[1], Debbie A Lawlor[1], Jon Heron[2], Abigail Fraser[1], Scott M Nelson[3], Kate Tilling[2]

[1] MRC Centre for Causal Analyses in Translational Epidemiology, School of Social and Community Medicine, University of Bristol, UK
[2] School of Social and Community Medicine, University of Bristol, UK
[3] Centre for Population and Health Sciences, University of Glasgow, UK

Background: Latent methods for sparse repeated categorical measurements are less developed than those for repeated continuous measures such as growth mixture modelling, or latent class analysis for repeated non-sparse categorical measures. We investigated latent classes of proteinuria occurrence during pregnancy, which was measured on a categorical scale with 99% of measurements in the null category of "nil/trace".

Methods and Results: Routine antenatal dipstick proteinuria assessments (median 12 per woman) were available for 11,716 women in the Avon Longitudinal Study of Parents and Children (ALSPAC) who had a live term birth without pre-eclampsia or pre-existing hypertension. We aggregated measurements to obtain the maximum proteinuria each woman experienced in eight periods of gestation: ≤12 weeks, 13-16 weeks, 17-20 weeks, 20-24 weeks, 25-28 weeks, 29-32 weeks, 33-36 weeks and 36+ weeks and fitted latent class models with 1 to 8 classes in Mplus using binary variables of "nil/trace" and "1+ or more" proteinuria. Due to the sparseness of the data this approach showed only two classes, representing women with low and high probabilities of experiencing proteinuria during pregnancy. We therefore fitted latent class models restricting to women who ever had proteinuria of 1+ or more (N=1,130) and categorised maximum proteinuria in each period as "nil/trace", "1+" and "2+ or more". Since there were particularly few occurrences of proteinuria in early pregnancy, bivariate residuals associated with the first three periods of gestation were large, violating the conditional independence assumption, so we combined these periods and used one measure for ≤20 weeks gestation. There were five latent classes of women who ever had proteinuria: women with proteinuria in early pregnancy only (≤20 weeks; N=101), and women with onset of proteinuria at 21-28 weeks (N=107), 29-32 weeks (N=137), 33-36 weeks (N=277) and 36 weeks gestation onwards (N=508). We used multinomial logistic regression with pseudo-class draws to investigate associations of maternal characteristics with odds of class membership, using the subgroup of women who never had proteinuria as the reference category. Maternal pre-pregnancy body mass index, age, parity, twin pregnancy and smoking during pregnancy were all associated with membership of at least one of the proteinuria latent classes.

Conclusions: We identified clinically meaningful classes representing the timing of onset of proteinuria and were able to distinguish characteristics associated with membership of each class. Our analysis suggests that care may be needed to avoid violating the conditional independence assumption in the presence of sparse data.

S7: MODELING THE HETEROGENEITY OF COGNITIVE DECLINE USING UNIVARIATE AND BIVARIATE GROWTH MODELS

Dr Graciela Muniz Terrera, MRC Biostatistics Unit, Institute of Public Health, University of Cambridge

Authors: Graciela Muniz Terrera, Louise Lafortune, Fiona Matthews, Carol Brayne

The general approach to studying older people's health has been to look at relationships among different dimensions of physical and cognitive health with the goal of identifying determinants and rates of decline and recovery, predicting mortality, and mapping service utilisation. Both conceptually and methodologically, the challenge has been to account for the fact that elderly populations are highly heterogeneous owing to the variability and interdependency in how health dimensions manifest themselves overtime. In this presentation we will focus on cognitive decline, which we know is determined by several domains of cognition and follows a non-linear process. We will explore the effect of jointly modelling age related cognitive decline of several cognitive domains when compared to univariate analyses. Then, we will examine if and how the shape of decline for individual dimensions differs - the underlying question being "Does it all go when it goes?" We used Mplus to fit the data from four interview waves (over a 9-year period; n= 2100) of the Cambridge City over 75 Cohort Study (CC75C) - a population based longitudinal study of aging.

S8: MULTIPLE IMPUTATION MULTIVARIATE MULTILEVEL MODEL OF WELL-BEING ACROSS EUROPE DURING THE CRISIS

Mr Adi Pierewan, University of Manchester

Most studies on happiness and health tend to separate them as a single outcome. This study investigates the individual and the country level covariation in happiness and health simultaneously as outcomes. Using data from the 2008 European Values Survey, we find that a large number of missing data appears in our variables. Therefore, we are in progress using multiple imputation, provided by Mplus using MCMC estimation, to overcome this problem by generating six imputed datasets. To analyse the data, we use multivariate multilevel model using both Mplus and MLwiN software. The results from these softwares are relatively similar. In preliminary test, without result from imputed data, we find that both at individual and country levels, happiness and health are positively correlated. We also find that in individual level: educated, married, and higher income people are positively associated with happiness and health. Conversely, unemployed and older people are negatively associated with happiness and health. In country level, we find that GDP and GDP growth are weakly and positively related with happiness and health; and Eastern Europe countries are weakly and negatively related with happiness and health.

S9: CLASS-BASED EFFECTS OF MEDIA VIOLENCE ON AGGRESSION: A MULTI LEVEL LATENT GROWTH MODEL APPROACH INVESTIGATING INDIVIDUAL AND GROUP DIFFERENCES

Mr Robert Busching, University of Potsdam

Authors: Robert Busching, Barbara Krahé, Ingrid Möller

Survey studies on aggression-related issues with school-aged children and adolescents are frequently conducted in classroom settings. However, class-based effects are often neglected while analyzing the data and therefore correlations among the examined constructs may be under- or overestimated. Multi-level modeling can help to solve this problem: Variables describing the group context can be included in the analysis, and differences between classrooms can be examined. In 78 classes 2064 adolescents' self-reports on habitual media violence exposure, aggression-enhancing normative beliefs, and perceived peer acceptance of aggression were assessed three times over a period of 36 months. Additionally, teacher ratings of students' aggression were obtained. Mean differences in media violence consumption and aggression were found between classes. Furthermore, the relationship between media violence exposure and the change in aggression over time differed between classes. In a three-level latent growth analysis, the teacher ratings were treated as an ordinal outcome variable. Results indicate that the individual and the class component of media violence exposure as well as their interaction at T1 predicted the slope of teacher-reported aggression. The next step will be the investigation of aggression-enhancing normative beliefs and peer acceptance of aggression as potential mediators using Multilevel SEM. The finding that classroom context is an important factor for analyzing the relationship between media violence and aggression is discussed in terms of its methodological and theoretical implications.

S10: Early life patterns of common infection: A latent class analysis

Mrs Sarah Fleming, University of Leeds

Authors: Sarah J Fleming [1], Graham R Law [1], Debbie A Lawlor [2,3], Patricia A McKinney [1]

[1] Centre for Epidemiology and Biostatistics, Leeds Institute of Genetics, Health and Therapeutics, University of Leeds, Leeds, UK
[2] Division of Epidemiology, Department of Social Medicine, University of Bristol, Bristol, UK
[3] Medical Research Council Centre for Causal Analyses in Translational Epidemiology, University of Bristol, Bristol, UK

Early life infection is an important factor implicated in the aetiology of many chronic diseases. Data on ten infectious symptoms were collected by parental questionnaire when children were 6 months old as part of the Avon Longitudinal Study of Parents and Children (ALSPAC), United Kingdom. A latent class analysis was used to identify patterns of infection and their relationship to five factors commonly used as proxies or previously associated with infection: sex, other children in the home, maternal smoking, breastfeeding and maternal education. A total of 10,032 singleton children were included in the analysis. Five classes were identified with differing infectious disease patterns and children were assigned to the class for which they had a highest probability of membership based on their infectious symptom profile: 'general infection' (n=1252, 12.5%), 'gastrointestinal' (n=1902, 19.0%), 'mild respiratory' (n=3560, 35.5%), 'Colds/ear ache' (n=462, 4.6%) and 'healthy' (n=2856, 28.5%). Females had a reduced risk of being in all infectious classes, other children in the home were associated with an increased risk of being in the 'general infection', 'mild respiratory' or 'ear ache' class. Breastfeeding reduced the risk of being in the 'general infection' and 'gastrointestinal' classes whereas maternal smoking increased the risk of membership. Higher maternal education was associated with an increased risk of being in the 'mild respiratory' group. Other children in the home had the greatest association with infectious class membership. Latent class analysis provided a flexible method of investigating the relationship between multiple symptoms and demographic and lifestyle factors.

S11: A LATENT CLASS APPROACH TO ADOLESCENT ALCOHOL PATTERNS: THE ALSPAC STUDY

Mr Roberto Melotti, University of Bristol

Objective: To investigate the patterns of alcohol use among adolescents of a birth cohort participating in the ALSPAC study (UK) and their associations with early measures of socioeconomic position (SEP).

Methods: 2475 boys and 2771 girls attended a clinical assessment at age 15. Measures of alcohol included age at first drink, drinking occasions in the past six months, drinking five or more drinks on any single day (binge drinking) and experiencing any alcohol problems in the past two years. We fitted a set of exploratory latent class analysis (LCA) models, among boys and girls as separate entities, to investigate their patterns of use in Mplus. Upon selection of the best model fit, we exported the resulting classes and their assignment probabilities for each observation to stata for further analyses. Socioeconomic variables (parental social class, maternal education, disposable income) and previous substance use data were then multiply imputed, either including or excluding the indicator variables for the LCA. Both multiple imputation procedures used the ice command in stata with the iweight option for the latent class assignment probabilities and also included the latent classes. The associations between current patterns of alcohol use and multiple SEP indicators were primarily assessed among complete cases by generalized ordered logistic regression. Results were then compared between this model and the analyses after imputation according to both procedures above.

Results: Patterns of alcohol use for boys and girls were highly consistent. Latent class analysis suggested four orderable groups in both genders: "non drinkers", "moderate drinkers", "occasional heavy drinkers" and "problematic drinkers". For each higher quintile of household income adolescents were progressively more likely to be in a higher drinking category (POR 1.06, 95%CI 1.01-1.11). "Problematic drinkers" were also more likely to report a higher number of alcohol related problems than any other group. Adjustment for previous substance use did not alter the associations with SEP. Results after multiple imputation were consistent with the complete cases analysis. The two imputation procedures yielded virtually equivalent results.

Conclusion: There were very similar patterns of alcohol use in boys and girls aged 15 in a British birth cohort. We found that boys and girls of higher income households were more likely to consume alcohol and this was consistent using different settings for the analysis. This might reflect the greater accessibility to alcohol in higher income households.

Biostatistics Group Seminar: Latent Variable Modelling

Intergenerational correlations in size at birth and the contribution of environmental factors: the Uppsala Birth Cohort Multigenerational Study, Sweden, 1915-2002
Dr Bianca De Stavola, London School of Hygiene and Tropical Medicine

Size at birth in parents and their children are known to be correlated, reflecting in part the influence of foetal and maternal genes. Socio-demographic factors, regarded as aspects of shared environment across generations, would also be expected to contribute, but evidence is limited. The aim of this study is to quantify the role of shared environment in explaining intergenerational correlations in birth weight and length, using data across three consecutive generations from the Uppsala Birth Cohort Multigenerational Study. This includes birth and socio-demographic data on 7,657 singletons, born in Uppsala 1915-1929 (G1), and their grandchildren (G3). Standard regression and biometrical models were used to study the correlations in size at birth of G1-G3 pairs. The data showed stronger correlations in maternal than paternal pairs for birth weight but not for birth length. These correlations were not reduced by adjustment for socio-demographic factors in regression models. In contrast, significant shared environment contributions to the intergenerational correlations were identified in biometrical models, averaging 14% for both birth measures. These models are a particular specification of structural equation models where latent variables are specified to capture the genetic and environmental factors shared by relatives. The results show that the two approaches lead to partly inconsistent results. Reasons of why this happens will be discussed and some general conclusions will be drawn.

Psychometric epidemiology, Rasch psychiatry and advances towards well-being
Dr Tim Croudace, University of Cambridge

Psychometric measurement models from educational testing have diversified over the years into health, quality of life and psychiatric research. Modern software developments make it possible for researchers to combine many useful model components in novel and flexible ways; the universe of models is large. However several distinct model types remain somewhat software bound in terms of application areas and communities. Recent experiences at the forefront of software capabilities and from interacting with UK psychometric communities will be summarised in the context of measurement challenges in mental health. Latent variable models discussed include latent class, trait and hybrid models. Applications relate to psychological distress and positive mental health and the notion of categorical versus dimensional spectrum from Rasch and general latent variable modelling perspective.

Interactions, effect moderation and random coefficients in latent variable models: identifying patterns of prenatal vulnerability
Prof Andrew Pickles, Institute of Psychiatry, KCL

Following the foetal origins hypothesis of Barker there is now wide & intense interest in developmental programming that can have both adaptive benefits and costs far into the organism's future. One of these programming mechanisms involves the HPA system, where it has been found that in rodents and other animals maternal behaviour can modify genetic predisposition to anxiety through methylation of the DNA - an epigenetic process. Other possibilities include the process of brain-stem maturation setting levels of autonomic response to stress (vagal tone and vagal withdrawal) which might also be modified by post-natal parent behaviour. We explore this possibility in humans using the Wirral Child Health and Development Study and the GLLAMM analysis framework. We estimate a model in which a maternal behaviour latent variable derived from 4 ordinal items by an Item-Response Theory measurement model, moderates the impact of both maternal depression at 32-weeks of pregnancy and poor intra-uterine growth, on 6-month infant vagal tone and on vagal withdrawal in response to a social stressor.